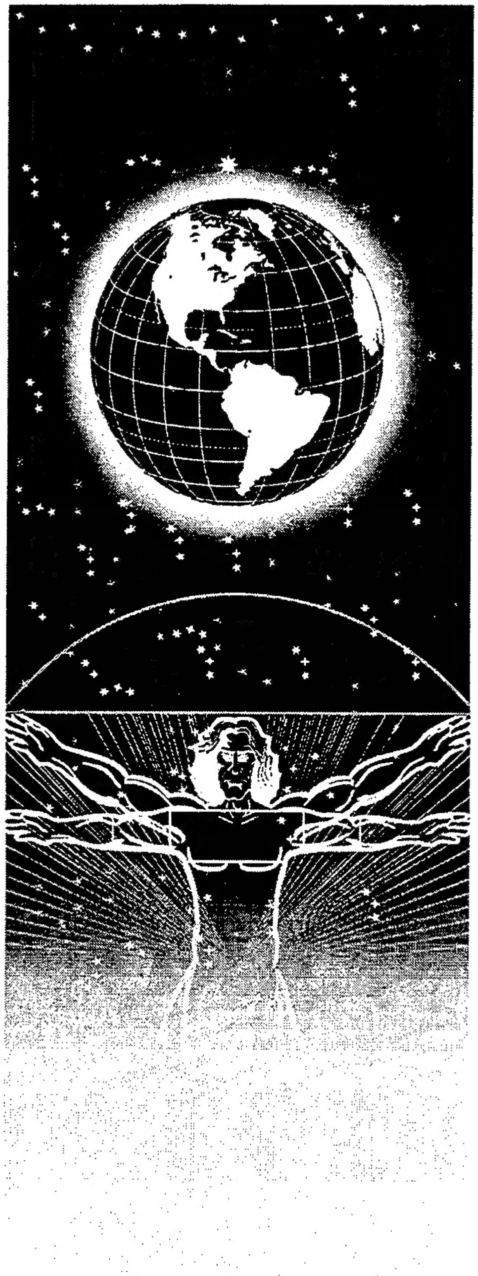
AL/HR-TP-1997-0033



UNITED STATES AIR FORCE
ARMSTRONG LABORATORY

MALE AND FEMALE CAUSAL MODELS OF PILOT SKILL ACQUISITION: A PRELIMINARY EVALUATION

Thomas R. Carretta

AIRCREW PERFORMANCE BRANCH
AIRCREW TRAINING RESEARCH DIVISION

Malcolm James Ree

COGNITION AND PERFORMANCE RESEARCH BRANCH

ARMSTRONG LABORATORY
HUMAN RESOURCES DIRECTORATE
7909 Lindbergh Drive
Brooks Air Force Base TX 78235-5352

September 1997

AIR FORCE MATERIEL COMMAND
ARMSTRONG LABORATORY
HUMAN RESOURCES DIRECTORATE
AIRCREW TRAINING RESEARCH DIVISION
6001 South Power Road, Building 558
Mesa AZ 85206-0904

NOTICES

Publication of this paper does not constitute approval or disapproval of the ideas or findings. It is published in the interest of scientific and technical information (STINFO) exchange.

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The Office of Public Affairs has reviewed this paper, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This paper has been reviewed and is approved for publication.

THOMAS R. CARRETTA
Project Scientist

DEE H. ANDREWS
Technical Director

LYNN A. CARROLL, Colonel, USAF
Chief, Aircrew Training Research Division

Notify AL/HRPP, 7909 Lindbergh Drive, Brooks AFB TX 78235-5352, if your address changes or if you no longer want to receive our technical reports. You may write/call the STINFO Office at DSN 240-3877 or Comm (210) 536-3877.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE
September 1997

3. REPORT TYPE AND DATES COVERED
Interim - January 1996 to April 1997

4. TITLE AND SUBTITLE

Male and Female Causal Models of Pilot Skill Acquisition: A Preliminary Evaluation

5. FUNDING NUMBERS

PE - 62205F
PR - 1123
TA - A1
WU - 01

6. AUTHOR(S)

Thomas R.Carretta
Malcolm J. Ree

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Armstrong Laboratory Human Resources Directorate
Aircrew Training Research Division
Aircrew Performance Branch, and Cognitive and Performance Research Division
7909 Lindbergh Drive
Brooks AFB TX 78235-5352

8. PERFORMING ORGANIZATION

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Armstrong Laboratory
Human Resources Directorate
Aircrew Training Research Division
6001 South Power Road, Building 561
Mesa AZ 85206-0904

10. SPONSORING/MONITORING

AL/HR-TP-1997-0033

11. SUPPLEMENTARY NOTES

Armstrong Laboratory Technical Monitor: Dr Thomas R. Carretta, (210) 536-3956
This paper has been submitted to the International Journal of Aviation Psychology.

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Based on a previous study, a causal model of acquisition of pilot job knowledge and flying skills was tested on separate samples of male and female students. Causal model parameters were estimated separately for each sample and, due to the small sample size for females, no between-groups statistical tests were conducted. The results are viewed as tentative because of the small sample of female students, however, the path coefficient parameter estimates are still useful. The model showed a direct influence of general cognitive ability on the acquisition of job knowledge and an indirect influence on the acquisition of flying skills. The direct and indirect influence of cognitive ability on flying skills was a little stronger for females than for males. Additionally, the path between prior job knowledge and flying performance was somewhat stronger for females than for males. Consistent with previous findings, the influence of early flying skills on later flying skills was very strong. No argument for a sex-separated training syllabus is supported.

14. SUBJECT TERMS
Acquisition, Causal model; Cognitive ability; Females; Flying performance; Flying skills; Job knowledge; Males; Path model; Pilots; Skills

15. NUMBER OF PAGES
20

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNLIMITED

CONTENTS

FIGURES

TABLES

PREFACE

This effort was conducted under Work Unit 1123A101, Pilot Selection and Classification Support, which is dedicated to research into the selection and classification of US Air Force aircrew personnel. The Laboratory Principal Investigator was Dr Thomas R. Carretta. The authors thank Mr Paul Rioux for his assistance in the development of the databases used in this research study.

Address all correspondence and requests for reprints to the first author at AL/HRAA, 7909 Lindbergh Drive, Brooks AFB, TX 78235-5352.

Send email to CARRETTA@ALHRM.BROOKS.AF.MIL.

MALE AND FEMALE CAUSAL MODELS OF PILOT SKILL ACQUISITION: A PRELIMINARY EVALUATION

INTRODUCTION

Historically, measures of general cognitive ability, *g*, and prior job knowledge have demonstrated consistent validity against pilot training performance (Carretta & Ree, 1994; Olea & Ree, 1994; Ree & Carretta, 1996). More recently, Ree, Carretta, and Teachout (1995) using latent variable path analysis, have demonstrated that *g* works through job knowledge to cause pilot performance. This finding--that *g* works through job knowledge to cause performance-- is consistent across studies of numerous jobs (Hunter, 1986). The current experiment evaluated a previously confirmed causal model of pilot training performance on separate male and female samples.

Male-Female Differences

Halpern (1992) argued the necessity of conducting research on sex differences noting that knowledge is preferable to ignorance. Differences between the sexes on mean score on ability tests have a long history. Tyler (1965) provides a useful overview as does Willerman (1979). In a meta-analysis, Hyde (1981) found the following median standardized mean differences on tests identified as measuring cognitive ability factors for men and women: .24 for verbal favoring women, .43 for quantitative favoring men, .45 for visual-spatial favoring men, and .51 for field articulation (defined as visual-analytic ability) favoring men. Burke (1995) observed that tests used in aviation selection are frequently those that favor men in mean score comparisons and called for the use of a compensatory model that balances the strengths of males and females.

Carretta (1997) examined mean score sex differences for the 16 tests used for United States Air Force (USAF) officer commissioning and pilot selection purposes. He found that large mean score differences between the sexes in officer commissioning applicant samples were substantially reduced among pilot trainees. Among the applicants, the standardized difference values favored the males for all 16 tests, although some were rather small. The mean standardized value was .44. After selection into pilot training, the standardized difference values were reduced, with a mean of .05.

Although groups may differ in means on tests, they may show similarity in the factors underlying those scores. Michael (1949); Humphreys and Taber (1973); Defries et al., (1974) studied factor similarity between ethnic groups and found few differences. Carretta and Ree (1995) and Ree and Carretta (1995) studied ethnic and sex group ability factor differences. For both ethnic and sex groups, they found a near identity of aptitude factor structure in both experiments with cross-group test loading correlations approaching 1.

Causal Models

Increasingly, path or causal models have been used to explain the relationships of variables in occupational settings. Hunter (1986) demonstrated the most general model relating ability, job knowledge, and job performance. He noted both a direct path from ability to job performance as well as an indirect path through job knowledge. His verified model showed ability leading to the acquisition of job knowledge which, in turn, led to job performance. Using cumulated meta-analyzed data, Hunter found a stronger direct path between ability and job performance for civilian versus military jobs. Ree et al. (1995) found a similar weak direct path in a military sample and, along with Hunter, speculated that the weak path is the result of the necessity to learn and apply myriad complex rules and procedures.

The Ree et al. (1995) model found significant causal paths relating ability (*g*), prior job knowledge (JK_p), sequentially acquired training job knowledge (JK_{T1}, JK_{T2}, and JK_{T3}) and work sample performance (WS_1 and WS_2). This model is shown in Figure 1.

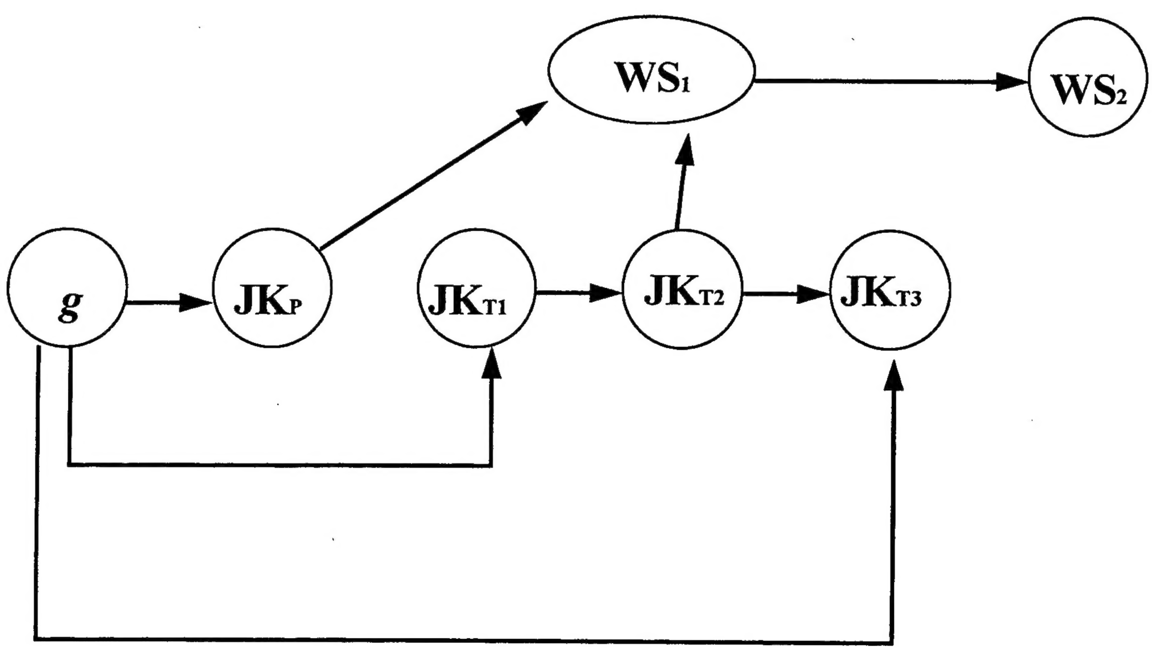


Figure 1. Hypothesized Causal Model for Sequential Training

There was a strong influence of *g* on the acquisition of all job knowledge. The early acquisition of job knowledge led to the later acquisition of job knowledge. Job knowledge showed a causal influence on early job performance as measured by flying work samples. Early job performance had a very strong causal influence on subsequent job performance.

Given the evidence of validity of *g* and job knowledge and the near identity of factor structure for the sexes, it is appropriate and informative to ask whether the same causal model would hold for each sex group. This experiment provides a preliminary answer to this question. If a different causal model were found for each sex group, this might be interpreted as justification for a separate training syllabus for men and women.

The results of this experiment must be interpreted with caution because of the small number of female pilots. A bigger sample would have been preferable, however, the total number of female pilots in the United States Air Force is very small and this sample represents a sizable portion of that total. Although the sample is small, it is presented so that it may be aggregated with other samples in future meta-analysis.

METHOD

Participants

The participants were 3,369 male and 59 female USAF officers who completed a 53-week undergraduate pilot training course between 1981 and 1993. They were predominantly white (96.8%), between about 22 and 27 years old, and had completed at least a baccalaureate degree from an accredited university or college. All had been selected for commissioning and undergraduate pilot training, in part, based on their scores on the Air Force Officer Qualifying Test (AFOQT) (Carretta & Ree, 1995; Skinner & Ree, 1987).

A selection board technique that rates applicants for admission to flying training is used by the USAF. Included are measures of academic achievement with a preference toward scientific majors, personal recommendations, medical fitness, and prior flying experience in some cases. These data are not retained in official archival files and were not available for this research effort.

Measures

g and Prior Job Knowledge

The measures of *g* and prior job knowledge were extracted from the AFOQT (Carretta & Ree, 1995, 1996). The AFOQT is based on a detailed taxonomy of test and item specifications that define the psychometric properties as well as the content of each test (Berger, Gupta, Berger, & Skinner, 1990; Gupta, Berger, Berger, & Skinner, 1989; Skinner & Ree, 1987).

The 16 tests that comprise the AFOQT provide measures of general cognitive ability (*g*), flying job knowledge, and four lower-order cognitive factors: verbal, quantitative, spatial, and perceptual speed (Carretta & Ree, 1996). In this experiment, verbal and quantitative tests-- the most universally accepted measures of general cognitive ability--were used to estimate *g*. The Instrument Comprehension and Aviation Information tests were used to assess prior job knowledge (JK_p). Provided below are descriptions of the tests grouped by content.

Verbal tests. Verbal Analogies (VA) measures the ability to recognize relationships between words and to reason. Reading Comprehension (RC) assesses the ability to understand written paragraphs. Word Knowledge (WK) provides a measure of verbal ability through the use of synonyms.

Quantitative tests. Arithmetic Reasoning (AR) measures the ability to understand arithmetic relationships stated as word problems. Data Interpretation (DI) assesses the ability to extract information from tables and charts. Math Knowledge (MK) requires the ability to use mathematical formulas, terms, and relationships to solve problems. Scale Reading (SR) measures the ability to extract information from scales and dials.

Prior job knowledge tests. Only two tests in the AFOQT measure specific job knowledge (Dye, Reck, & McDaniel., 1993; Olea & Ree, 1994). Instrument Comprehension (IC) assesses the ability to determine the position and orientation in three-dimensional space of an aircraft in flight based on illustrations of flight instruments. Aviation Information (AI) measures knowledge of general aviation concepts, principles, and terminology.

Pilot Academic and Flying Grades

Pilot academic grades. Academic indicators measured student pilots' performance on written tests of flying theory, procedures, and aircraft-unique systems (i.e., hydraulics, instruments, electronics, etc.) learned during training. On each academic test, each student received a percent correct score. There were 11 end-of-course tests (A1 through A11) that were divided into three groups to represent early (A1 to A4), middle (A5 to A8), and late (A9 to A11) training. Early and middle classroom training were relevant to flying the subsonic primary training aircraft (T-37). Early classroom training included courses in T-37 systems, T-37 aerodynamics, aerospace physiology/human factors, and flying fundamentals. Middle classroom training provided courses relevant to flight in general and to flying the primary aircraft. Included were T-37 instruments I and II, T-37 navigation, and T-37 mission planning. Late classroom training was relevant to the supersonic advanced training aircraft (T-38) including applied aerodynamics, T-38 systems operations, and T-38 flight planning.

Flying work samples. There are two general categories of training flights in which students accumulate about 190 flying hours. On routine daily flights, the student pilot learns and practices under the watchful eye of an instructor pilot. After the prescribed ordinary daily flights, work-sample tests called "check flights" are rated by check flight pilots. Check flight pilots do not rate students with whom they have flown on daily flights to eliminate potential bias due to familiarity.

Three check flights in the primary aircraft (CF1 to CF3) and three in the advanced aircraft (CF4 to CF6) are completed by student pilots during training. In the primary aircraft, students must (a) demonstrate the ability to fly to a geographical location, perform aerial maneuvers, and return to execute successful landings; (b) conduct airborne activities within precise geographical and altitude limits; and (c) use instruments with an emphasis on landing approaches.

All activities must be accomplished more rapidly in the advanced training aircraft because it is much faster than the primary training aircraft. This makes even familiar maneuvers more difficult.

The check flights for instruments and round trips to geographical areas are similar to the check flights in the primary aircraft. The difficult formation check flight is added in which the wings of multiple aircraft are as close as three feet at speeds of 400 knots. See Duke and Ree (1996) for a more complete description of check flights in the advanced aircraft.

Each check flight score was a weighted average of ratings of several flying maneuvers and procedures. These maneuvers, procedures, and scoring weights are prescribed by the Air Force in training regulations. The student pilot receives points for each procedure. Example procedures are: make proper radio calls during flight, retract landing gear at specified speed, or perform loop within specified parameters (e.g. maneuver entry altitude and engine power settings). Like academic grades, check flight grades were percentage scores.

The sequential pilot training was structured as follows. In the classroom, theory and general background were taught. This was followed by application in the aircraft. Classroom training for the primary aircraft began before check flight work samples. The ultimate check flight work sample in the primary training aircraft was completed after the last classroom instruction in middle training (A5 to A8). After check flights in the primary aircraft, classroom instruction on the advanced aircraft began. This was followed by advanced aircraft check flight work samples. The last advanced aircraft check flight work sample occurred after all classroom training was completed.

Procedures

The current experiment investigated the causal role of *g* and prior job knowledge for both men and women in flying training. Included were measures of *g* and job knowledge acquired prior to training, sequentially-ordered blocks of classroom training, and hands-on flying work sample performance measures.

The participants constituted a censored, range-restricted sample because they had been selected, at least in part, on the basis of the scores of the test battery that yielded the estimates of *g* and prior job knowledge. To correct the poor statistical estimates of the correlations among variables found in range restricted samples (Thorndike, 1949), we used the multivariate method of Lawley (1943; see also Ree, Carretta, Earles, & Albert, 1994). Male samples were corrected to a group of male applicants and females were corrected to a group of female applicants. Corrected matrices were used in all structural equation analyses.

The structural models (Bentler & Weeks, 1980) were estimated using maximum likelihood procedures as implemented in version 4.02 of the EQS program. This program corrects for unreliability using estimation procedures in the same fashion as LISREL and other structural modeling programs. The estimated reliabilities can either be provided as starting values or they can be estimated directly from the data as was done here.

First we fit the measurement models and then the path models as established in previous research. We reported the path coefficients as standardized regression coefficients (Cohen & Cohen, 1983) because the scales of measurement of the variables are not well known or intrinsically meaningful. These standardized path coefficients should be interpreted as indicating that a one standard deviation change in an independent variable leads to a change in the dependent variable equal to the magnitude of the coefficient. For example, if the path coefficient were .75, a one

standard deviation change in the independent variable would yield a .75 standard deviation change in the dependent variable.

Path models based on Ree et al. (1995) with only the statistically significant links were estimated for separate male and female samples.

RESULTS

The means and standard deviations for the variables both in observed form and after range restriction correction are presented in Table A1. Tables A2 and A3 present the correlation matrices for the male and female samples, both observed and corrected for range restriction.

Table 1.
Correlations Between Factors in the Causal Model

Factor	*g*	JK_P	JK_{T1}	JK_{T2}	JK_{T3}	WS_1	WS_2
Males							
g	1.00						
JK_P	.63	1.00					
JK_{T1}	.62	.42	1.00				
JK_{T2}	.55	.29	.87	1.00			
JK_{T3}	.59	.30	.85	.94	1.00		
WS_1	.32	.29	.44	.56	.54	1.00	
WS_2	.37	.36	.43	.54	.55	.91	1.00
Females							
g	1.00						
JK_P	.80	1.00					
JK_{T1}	.76	.67	1.00				
JK_{T2}	.59	.36	.85	1.00			
JK_{T3}	.84	.68	1.00	1.00	1.00		
WS_1	.71	.62	.32	.54	.73	1.00	
WS_2	.81	.50	.60	.82	.50	1.00	1.00

Note. *g* = general cognitive ability; JK_P = prior job knowledge; JK_{T1} = job knowledge acquired during training (measure 1); JK_{T2} = job knowledge acquired during training (measure 2); JK_{T3} = job knowledge acquired during training (measure 3); WS_1 = flying training work sample (measure 1); and WS_2 = flying training work sample (measure 2).

The intercorrelations of the factors as estimated from the corrected data for each sample are presented in Table 1. The variance accounted for in each dependent variable is presented in Table 2 and the structural coefficients are shown in Figure 2.

Table 2.
Variance Accounted for (R^2) in the Dependent Variables

	Males (n = 3,369)	Females (n = 59)
JK_P	.395	.658
JK_{T1}	.396	.599
JK_{T2}	.742	.628
JK_{T3}	.935	1.000
WS_1	.333	.570
WS_2	1.000	1.000

Note. JKp = prior job knowledge; JK_{T1} = job knowledge acquired during training (measure 1); JK_{T2} = job knowledge acquired during training (measure 2); JK_{T3} = job knowledge acquired during training (measure 3); WS_1 = flying training work sample (measure 1); and WS_2 = flying training work sample (measure 2).

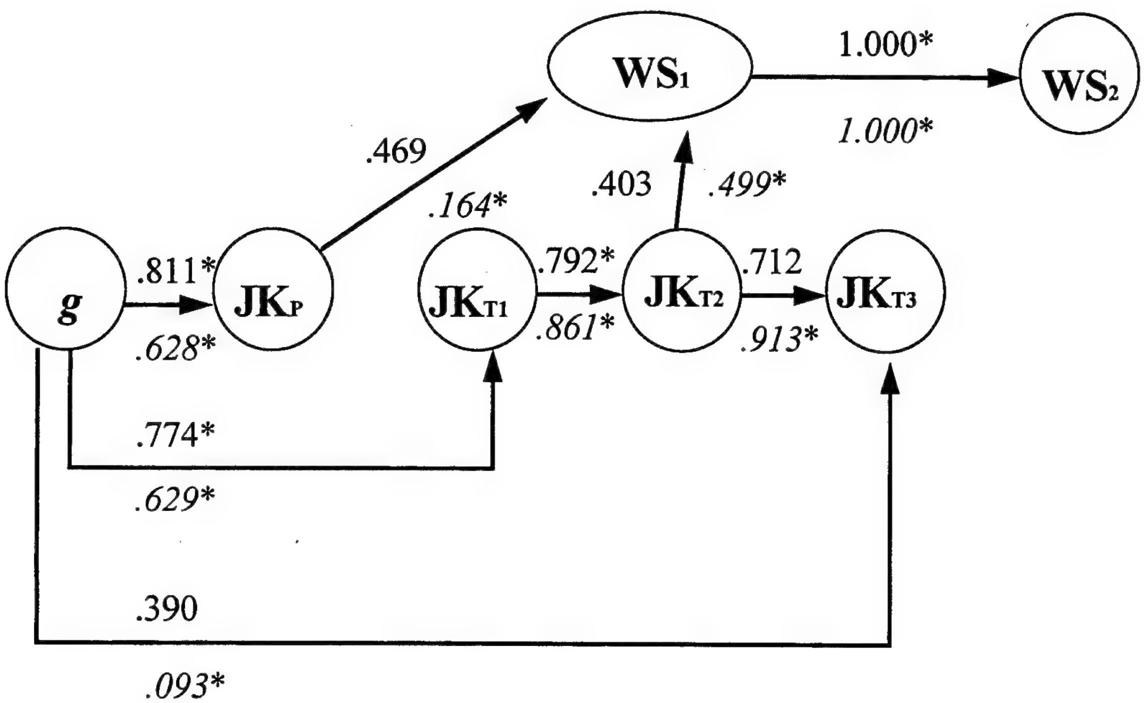


Figure 2. Causal Model for Sequential Training (Independent Male and Female Samples)

Note. (1) Male (n = 3,369) and female (n = 59) causal models were estimated independently. (2) All path coefficients were significant at $p < .05$. Those marked with an * were significant at $p < .01$.

DISCUSSION

Group mean differences on the verbal and quantitative tests, measures of *g*, favored women. The opposite was true for the tests of prior job knowledge. The average standardized differences (*d*) on the verbal and quantitative tests were -.33 and .59 for the prior job knowledge. Each sex group brings different strengths to the training situation.

The means, standard deviations, and correlations after correction for range restriction represent the best statistical estimates. As would be expected, the corrected standard deviations increased and the corrected means decreased. The corrected correlations behaved in accordance with Lawley's (1943) theorem. For the large sample of males, positive manifold was observed. In the sample of females, the correlations were mostly positive. The reason for the lack of total positive manifold in the female sample cannot be known from these data, but variability due to small sample size is a reasonable explanation. The correlations for the factors show positive manifold for both the male and female samples.

The structural coefficients for the models for each group estimated independently showed general similarity, but with some differences. The causal effect of *g* on prior job knowledge was strong for both sexes. This was also true of the causal path from *g* to job knowledge acquired during training. A notable exception was the much greater influence for females than for males of g on JK_{T3} and WS_2. The total causal influence of *g* on JK_{T3} was .826 for females and .587 for males. Similarly, the total causal influence of *g* on WS_2 was .627 for females and .373 for males. However, it should be noted that the variance accounted for on JK_{T3} and WS_2 was about the same for both males and females. This can be interpreted as showing that the antecedents have about the same cumulative effect.

The causal influence for prior job knowledge on work sample performance appeared weaker for males than for females. The causal influence of job knowledge acquired during training on subsequent job knowledge and work sample performance was stronger for males than for females as shown by the coefficients between JK_{T1} and JK_{T2} and between JK_{T2} and JK_{T3}.

Because of the small differences between men and women in the causal paths from *g* to JK_{T1} to JK_{T2} to JK_{T3}, it appears that the dependence on ability and job knowledge for the acquisition of later job knowledge is similar for both groups. Further, because the variance accounted for in JK_{T3} and WS_2, the two end-of-training dependent variables, was about equal for men and women, any argument for a sex-separated training syllabus is not supported.

REFERENCES

Bentler, P. M., & Weeks, D. G. (1980). Linear structural equations with latent variables. *Psychometrika*, *45*, 289-308.

Berger, F. R., Gupta, W. B., Berger, R. M., & Skinner, J. (1990). *Air Force Officer Qualifying Test (AFOQT) form P manual* (AFHRL-TR-89-56). Brooks Air Force Base, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Burke, E. (1995). Male-female differences on aviation selection tests: Their implications for research and practice. In N. Johnston, R. Fuller, and N. McDonald (Eds.), *Aviation Psychology: Training and Selection* (pp. 188-193). Aldershot, UK: Avebury Aviation.

Carretta, T. R. (1997). Group differences on U. S. Air Force pilot selection tests. *International Journal of Selection and Assessment*, *5*, 115-127.

Carretta, T. R., & Ree, M. J. (1994). Pilot candidate selection method: Sources of validity. *The International Journal of Aviation Psychology*, *4*, 103-117.

Carretta, T. R., & Ree, M. J. (1995). Near identity of cognitive structure in sex and ethnic groups. *Personality and Individual Differences*, *19*, 149-155.

Carretta, T. R., & Ree, M. J. (1996). Factor structure of the Air Force Officer Qualifying Test: Analysis and comparison. *Military Psychology*, *8*, 29-42.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analyses for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Defries, J. C., Vandenberg, S. G., McClearn, G. E., Kuse, A. R., Wilson, J. R., Ashton, G. C., & Johnson, R. C. (1974). Near identity of cognitive structure in two ethnic groups. *Science*, *183*, 338-339.

Duke, A. P., & Ree, M. J. (1996). Better candidates fly fewer training hours: Another time testing pays off. *International Journal of Selection and Assessment*, *4*, 115-121.

Dye, D. A., Reck, M., & McDaniel, M. (1993). The validity of job knowledge measures. *International Journal of Selection and Assessment*, *1*, 153-162.

Gupta, W. B., Berger, F. R., Berger, R. M., & Skinner, J. (1989). *Air Force Officer Qualifying Test (AFOQT): Development of an item bank* (AFHRL-TP-89-33). Brooks Air Force Base, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Halpern, D. F. (1992). *Sex differences in cognitive abilities* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hunter, J. E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior*, *29*, 340-362.

Humphreys, L. G., & Taber, T. (1973). Ability factors as a function of advantaged and disadvantaged groups. *Journal of Educational Measurement*, *10*, 107-115.

Hyde, J. S. (1981). How large are cognitive gender differences? A meta-analysis using ω^2 and *d*. *American Psychologist*, *36*, 892-901.

Lawley, D. N. (1943). A note on Karl Pearson's selection formulae. *Proceedings of the Royal Society of Edinburgh*, *62* (Section A, Pt. 1), 28-30.

Michael, W. B. (1949). Factor analysis of tests and criteria: A comparative study of two AAF pilot populations. *Psychological Monographs*, *63*, 55-84.

Olea, M. M., & Ree, M. J. (1994). Predicting pilot and navigator criteria: Not much more than *g*. *Journal of Applied Psychology*, *79*, 845-851.

Ree, M. J., & Carretta, T. R. (1995). Group differences in aptitude factor structure on the ASVAB. *Educational and Psychological Measurement*, *55*, 268-277.

Ree, M. J., & Carretta, T. R. (1996). Central role of *g* in military pilot selection. *The International Journal of Aviation Psychology*, *6*, 111-123.

Ree, M. J., Carretta, T. R., Earles, J. A., & Albert, W. (1994). Sign changes when correcting for range restriction: A note on Pearson's and Lawley's selection formulas. *Journal of Applied Psychology*, *79*, 298-301.

Ree, M. J., Carretta, T. R., & Teachout, M. S. (1995). Role of ability and prior job knowledge in complex training performance. *Journal of Applied Psychology*, *80*, 721-730.

Skinner, J., & Ree, M. J. (1987). *Air Force Officer Qualifying Test: Item and factor analysis of form O* (AFHRL-TR-86-68, AD A184 975). Brooks Air Force Base, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Thorndike, R. L. (1949). *Personnel selection*. NY: Wiley.

Tyler, L. (1965). *The psychology of human differences* (3rd ed.). NY: Appleton-Century-Crofts.

Willerman, L. (1979). *The psychology of individual and group differences*. San Francisco: W. H. Freeman.

APPENDIX A

Means, Standard Deviations, and Correlation Matrices for Male and Female Pilot Trainees

Table A1.
Means and Standard Deviations for Tests, Academic Grades, and Check Flight Grades

	Observed				Corrected for Range Restriction			
	Males (n = 3,369)		Females (n = 59)		Males (n = 3,369)		Females (n = 59)	
Score	Mean	SD	Mean	SD	Mean	SD	Mean	SD
VA	15.27	3.36	16.76	2.85	14.25	4.34	14.13	4.73
AR	13.53	4.12	13.95	3.42	12.52	5.00	10.29	4.81
RC	17.39	4.73	20.34	3.78	15.80	5.62	15.24	5.87
DI	13.51	3.91	14.19	3.83	12.89	4.66	11.49	4.52
WK	13.91	5.16	17.00	5.02	13.86	5.73	13.35	5.94
MK	17.98	4.63	19.68	4.36	15.10	5.91	13.34	5.96
SR	24.23	5.55	24.44	5.48	21.55	6.61	18.28	6.64
IC	13.71	4.21	11.02	4.97	10.48	5.04	6.68	3.85
AI	11.78	4.25	9.42	4.12	9.11	4.13	6.05	2.79
A1	97.46	3.08	97.09	4.44	96.81	3.15	92.80	5.44
A2	97.17	3.33	97.62	3.09	96.68	3.38	95.11	3.49
A3	97.04	3.36	97.03	3.49	96.62	3.40	94.28	4.32
A4	98.07	3.28	97.50	3.26	97.37	3.36	97.44	3.32
A5	95.97	4.80	96.83	4.35	95.24	4.89	95.47	4.51
A6	95.17	5.33	95.70	4.52	94.52	5.38	93.97	5.45
A7	94.75	5.37	96.07	4.08	94.07	5.46	96.02	4.25
A8	95.86	4.56	97.23	3.95	95.34	4.61	93.89	4.44
A9	97.36	3.32	97.61	3.37	96.89	3.37	94.99	3.52
A10	97.29	3.63	97.62	3.70	96.80	3.69	96.74	3.96
A11	96.82	3.70	97.08	3.04	96.24	3.76	94.87	3.39
CF1	86.57	7.57	87.90	6.86	85.53	7.62	84.84	8.26
CF2	90.64	5.76	91.68	3.50	89.94	5.80	89.68	4.03
CF3	93.56	4.89	95.32	3.57	92.83	4.94	93.13	3.78
CF4	91.20	5.72	91.52	4.97	90.34	5.77	91.47	5.20
CF5	92.66	4.67	92.22	4.23	92.15	4.69	91.66	4.69
CF6	93.82	4.73	93.75	4.37	93.14	4.78	89.54	5.53

Notes. (1) The 9 AFOQT tests are: VA = Verbal Analogies, AR = Arithmetic Reasoning, RC = Reading Comprehension, DI = Data Interpretation, WK = Word Knowledge, MK = Mathematics Knowledge, SR = Scale Reading, IC = Instrument Comprehension, and AI = Aviation Information. The 11 flying training academic grades are A1 through A11. The 6 flying training check flight grades are CF1 through CF6.
(2) Means and standard deviations were corrected for range restriction (Lawley, 1943).

Table A2.
Correlation Matrix (Males, n = 3,369)

Score	VA	AR	RC	DI	WK	MK	SR	IC	AI	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	CF1	CF2	CF3	CF4	CF5	CF6
VA	100	48	57	47	60	39	33	28	18	16	17	14	15	12	08	11	11	14	13	11	01	05	06	01	03	05
AR	61	100	44	63	43	57	55	30	13	17	14	17	16	15	13	18	14	14	19	15	07	08	09	06	04	08
RC	66	49	100	43	69	38	32	25	22	17	16	14	17	10	09	12	12	15	15	12	00	03	06	02	01	06
DI	60	73	47	100	44	43	51	35	22	17	12	14	15	12	12	19	14	13	14	16	08	10	10	08	06	10
WK	66	45	73	46	100	32	29	28	29	13	15	12	12	06	07	10	10	11	11	08	00	02	03	-02	00	02
MK	52	68	46	55	37	100	42	19	01	18	15	15	18	16	12	15	12	16	22	18	05	07	12	07	05	12
SR	47	65	41	60	34	58	100	36	18	14	10	12	12	15	11	13	14	11	15	15	09	11	13	10	06	13
IC	42	41	32	45	28	37	48	100	45	12	09	09	11	06	05	07	06	06	05	07	10	09	08	10	09	08
AI	29	21	35	24	31	19	27	48	100	13	10	08	11	03	07	10	08	07	03	05	08	05	04	03	02	01
A1	23	24	23	23	17	26	22	19	19	100	19	18	23	20	17	14	12	19	19	14	07	08	12	07	04	08
A2	24	20	22	18	19	21	17	15	15	22	100	15	21	19	17	18	17	16	15	17	05	06	07	06	05	08
A3	20	23	18	19	16	22	18	14	12	21	17	100	17	21	19	17	15	19	15	14	07	06	14	05	05	06
A4	23	23	23	21	16	26	20	19	17	26	24	20	100	12	20	19	13	19	17	12	10	08	10	10	05	07
A5	19	23	16	19	10	24	22	13	09	23	22	24	16	100	21	21	24	24	18	17	14	11	16	08	09	12
A6	13	19	13	17	10	18	17	10	10	20	19	21	22	23	100	18	19	23	17	18	13	14	12	10	08	12
A7	18	25	16	25	13	22	20	13	13	17	20	19	22	24	20	100	20	23	14	20	11	10	10	09	08	10
A8	17	20	16	19	13	18	20	11	11	15	19	18	16	26	21	22	100	22	17	18	10	12	10	08	06	07
A9	21	20	20	18	16	22	18	12	12	22	19	22	22	26	25	25	24	100	21	20	10	11	11	10	05	09
A10	19	25	20	20	15	28	22	11	08	23	17	17	20	21	20	17	19	23	100	20	12	13	11	14	08	14
A11	18	22	17	22	12	25	22	14	10	18	20	16	15	20	20	23	20	23	23	100	12	11	13	09	05	11
CF1	05	11	02	11	00	10	14	14	10	09	06	09	12	16	14	12	11	12	14	14	100	25	25	21	21	17
CF2	10	13	07	14	04	12	15	13	08	11	08	08	10	13	15	12	13	13	15	13	26	100	28	19	19	17
CF3	11	15	10	15	05	18	18	13	08	15	09	16	13	18	14	12	12	13	14	15	26	29	100	19	19	19
CF4	06	11	04	12	00	12	15	15	07	09	08	07	12	10	12	11	10	11	15	11	22	20	20	100	19	16
CF5	07	08	03	10	01	09	10	14	04	06	06	06	07	11	09	10	07	06	09	07	22	20	20	20	100	15
CF6	11	14	10	15	05	18	18	13	05	11	10	09	10	15	14	12	09	12	16	14	19	19	21	18	16	100

Notes. (1) Correlations above the diagonal are observed data. Those below the diagonal were corrected for range restriction (Lawley, 1943).
(2) The 9 AFOQT tests Are: VA = Verbal Analogies, AR = Arithmetic Reasoning, RC = Reading Comprehension, DI = Data Interpretation, WK = Word Knowledge, MK = Mathematics Knowledge, SR = Scale Reading, IC = Instrument Comprehension, and AI = Aviation Information. The 11 flying training academic grades are A1 through A11. The 6 flying training check flight grades are CF1 through CF6.

13

Table A3.
Correlation Matrix (Females, n = 59)

Score	VA	AR	RC	DI	WK	MK	SR	IC	AI	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	CF1	CF2	CF3	CF4	CF5	CF6
VA	100	39	59	43	64	44	42	14	-07	27	23	16	-03	11	34	07	25	10	05	12	00	11	09	01	22	13
AR	64	100	38	59	39	46	39	15	-10	18	27	18	-07	23	18	-11	30	01	30	24	27	07	14	-14	24	00
RC	72	53	100	37	63	32	19	04	-06	27	30	18	-02	05	23	-11	25	17	-05	19	-03	09	-04	00	00	09
DI	60	70	47	100	40	48	55	19	-11	16	05	-05	-07	-05	03	06	04	-03	31	15	26	19	07	-10	16	02
WK	69	47	75	45	100	42	29	-05	-23	22	22	13	00	05	15	-07	24	07	-08	04	-02	-03	-10	08	03	05
MK	55	69	48	55	37	100	37	-06	-17	40	21	39	01	16	11	01	26	20	16	16	12	02	-05	-14	11	28
SR	55	66	47	61	38	61	100	37	01	11	13	02	-10	-02	10	-01	16	-03	14	09	39	40	38	02	19	13
IC	40	43	33	44	28	38	48	100	49	07	10	-03	03	-23	00	00	-03	16	-04	-05	07	28	38	07	00	05
AI	31	26	35	24	31	23	29	33	100	10	25	13	14	-02	01	32	04	34	00	10	-05	20	26	17	-11	-05
A1	47	43	47	33	34	61	38	31	32	100	36	33	24	21	40	13	20	36	25	26	-02	31	-03	-03	04	19
A2	40	51	45	30	39	44	41	24	39	50	100	30	26	13	47	13	46	33	15	06	-05	14	-10	05	15	04
A3	35	43	36	15	27	61	29	19	27	51	42	100	29	15	11	00	01	37	-05	00	-15	00	-01	10	07	13
A4	-08	-11	-07	-06	01	-02	-14	-03	04	12	15	20	100	16	24	-04	02	22	14	07	-13	04	-10	-04	-10	19
A5	31	39	20	13	17	33	29	00	15	34	26	27	10	100	29	-07	34	00	12	12	-08	07	-06	-19	32	07
A6	37	37	30	18	18	27	33	16	16	52	56	25	08	37	100	31	33	31	41	14	00	29	-10	-02	17	14
A7	06	00	-09	13	-06	00	01	-04	21	15	13	-02	-13	-03	40	100	-09	35	25	11	27	14	08	30	09	-02
A8	38	46	41	23	36	39	48	17	23	32	54	11	-06	39	37	-08	100	19	22	19	-13	13	00	-18	10	10
A9	19	22	31	11	21	41	17	21	31	48	41	47	14	09	39	27	25	100	20	01	-14	10	02	09	-26	20
A10	21	39	03	46	-03	29	33	11	09	34	26	05	07	18	49	34	28	25	100	16	19	10	09	-16	-07	21
A11	40	39	39	36	25	39	44	17	30	43	23	15	00	20	27	13	32	13	24	100	09	14	15	09	14	20
CF1	06	36	09	23	-07	21	45	13	06	12	13	-03	-21	07	28	35	11	01	31	23	100	20	24	23	10	09
CF2	03	18	12	20	-03	14	45	21	15	33	23	03	-05	16	43	20	25	15	22	26	41	100	30	04	35	06
CF3	38	35	27	28	18	23	51	31	25	14	02	09	-16	12	03	03	23	05	20	35	27	25	100	18	17	11
CF4	03	-16	05	-09	17	-22	-05	-05	12	-11	-04	-04	-04	-19	-10	20	-13	-01	-19	08	11	-05	15	100	01	00
CF5	36	35	06	32	06	23	35	17	05	18	25	13	-13	38	29	22	19	-15	09	26	20	40	29	-01	100	05
CF6	31	15	31	13	14	43	37	25	13	35	16	20	09	16	21	00	28	29	24	39	21	18	33	01	15	100

Notes. (1) Correlations above the diagonal are observed data. Those below the diagonal were corrected for range restriction (Lawley, 1943).
(2) The 9 AFOQT tests Are: VA = Verbal Analogies, AR = Arithmetic Reasoning, RC = Reading Comprehension, DI = Data Interpretation, WK = Word Knowledge, MK = Mathematics Knowledge, SR = Scale Reading, IC = Instrument Comprehension, and AI = Aviation Information. The 11 flying training academic grades are A1 through A11. The 6 flying training check flight grades are CF1 through CF6.

14